METHODOLOGY AND APPARATUS FOR SOLVING LOCKUP CONDITIONS WHILE TRUNKING IN FIBRE CHANNEL SWITCHED ARBITRATED LOOP SYSTEMS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001]   The present patent application is related to United States patent application no. 10/617,149 (the '149 application), entitled "Methods And Apparatus For Trunking In Fibre Channel Arbitrated Loop Systems" and filed on July 10, 2003, the contents of which are incorporated herein by reference in their entirety. The '149 application claims priority to United States patent application no. 10/612,753 filed on July 1, 2003, which claims priority to United States patent application no. 60/393,164 filed on July 2, 2002.

[0002]   The present patent application is also related to United States patent application no. 10/616,862 (the '862 application), entitled "Methods And Apparatus For Device Access Fairness In Fibre Channel Arbitrated Loop Systems" and filed on July 10, 2003, the contents of which are incorporated herein by reference in their entirety.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0003]   Not applicable.

REFERENCE TO A COMPACT DISK APPENDIX

[0004]   Not applicable.

BACKGROUND

[0005]   The '149 application (cited above in the RELATED APPLICATIONS section) describes employing trunking in a Fibre Channel network topology. Fibre Channel is a American National Standards Institute (ANSI) set of standards which describes a high performance serial transmission protocol which supports higher level storage and networking protocols.

[0006]   The '149 application describes, in the context of a Fibre Channel topology, trunking capabilities for back-end storage array designs. For example, the '149 application

describes, with reference to Fig. 27 of the '149 application, multiple duplicate cascades between Loop Switches in switching mode, to increase throughput between adjacent Loop Switches. Fig. 27 of the '149 application is reproduced herein as Fig. 1.

[0007] The duplicate cascades can be simultaneously activated without creating an invalid loop topology. Furthermore, load balancing between the cascades on an initiator basis is supported. In a system with two initiators and a primary/duplicate cascade pair, each initiator can have a logical cascade chain dedicated to it providing approximately 2x the throughput of a single cascade system.

[0008] Referring to Fig. 1 as an example, initiator HBA1 1801 has a full bandwidth path 1813, 1829, 1831 thru the string of SBODs 1803, 1804, 1805. Initiator HBA2 1802 also has a full bandwidth path 1814, 1830, 1832 thru the string of SBODs 1803, 1804, 1805. Simultaneous communication between HBA1 and a disk in an SBOD and HBA2 and a disk in an SBOD can occur. For example, HBA1 can communicate with Disk 1 1817 in SBOD 1803 using the path 1813, 1817 at the same time HBA2 1802 can communicate with Disk 16 1819 in SBOD 1803 using the path 1814, 1819. The number of duplicate cascades in a trunk is not limited by the hardware. A trunk group could be defined as 21 trunks in a 22 port ASIC if so desired (1 port must not be assigned to the trunk to provide the other side of the connection). If more initiators than cascades are added, throughput is affected based on the relative traffic assigned to each trunk within a group.

[0009] Load balancing among trunks is also discussed in the '149 application. Broadly speaking, most trunking implementations operate in either an autonomous or dynamic mode to balance traffic among trunks. Autonomous methods of load balancing generally assign trunks statically (for example, using round robin assignments). Dynamic load balancing actively monitors traffic on trunks and reassigns traffic flows to balance loads.

[0010] It is desired to minimize lockup conditions when initiating connections on the trunks.

BRIEF SUMMARY

[0011] Lockup conditions are solved while trunking in Fibre Channel switched arbitrated loop systems. Within a particular switch, a particular combination of pending OPN

2

conditions is detected, indicating a lockup condition. At least one of the detected pending OPN conditions is closed, which alleviates the lockup condition.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] FIG. 1 illustrates a trunking configuration in a Fibre Channel networking topology ;

[0013] FIG. 2, in conjunction with Table 1, illustrates a lockup condition occurring where multiple trunks are utilized in a switched arbitrated loop system;

[0014] FIG. 3 illustrates a configuration in which a lockup condition occurs in a configuration where switches are connected together by string ports.

[0015] FIG. 4 broadly illustrates a method to detect and alleviate a lockup condition.

DETAILED DESCRIPTION

[0016] Fig. 2 illustrates a lockup condition occurring in a situation where multiple trunks are utilized in a switched arbitrated loop system such as is described in the Background with reference to Fig. 1. In particular, with such a system, in some circumstances, traffic may stop until a disk or HBA times out and closes a connection. This can degrade performance to a level of unacceptability. One example of the problem is described with reference to Fig. 2.

[0017] However, before discussing the lockup condition with reference to Fig. 2, several terms are explained.

a) Pending Open: Assume that HBA#2 wants to connect to D1. First HBA#2 ARB's for the port it is connected to and SW#1 will pass it's ARB back to HBA#2. HBA#2 now has won arbitration and now can send its OPN, which is then received by SW#1. SW#1 now has a Pending Open on the port to which HBA#2 is physically connected.

b) Connection: Once there is a pending open, the router in SW#1 performs a lookup to determine if the destination port of the OPN is busy or free. If the destination port is free, then the router connects the switch matrix of SW#1 to the destination port. It is now said there is a Connection between HBA#2 and the OPN's destination port.

[0018]   We now turn to discussing a particular example of a lockup condition with reference to Fig. 2 and Table 1. With reference to Table 1, assume that HBA#1 (source) wants to connect to D3 (destination), and HBA#2 wants to connect to D1. In addition, D2 wants to connect to HBA#1 and D4 wants to connect to HBA#2. All of these devices are attempting to make the connection.

Source	Destination	Trunk	SW1	SW2	SW3
HBA#1	D3	T#4	Connected		Pending
HBA#2	D1	T#2	Connected	Pending	
D2	HBA#1	T#1	Pending	Connected	
D4	HBA#2	T#3	Pending		Connected

TABLE 1

[0019]   Furthermore, while the state of SW3 for the D4 to HBA#2 attempted connection is "connected," the state of SW1 for this attempted connection is "pending." This is a result of a previous attempted connection from HBA#2 to D1. That is, as can be seen in the row entry for HBA#2 to D1 in Table 1, the state of SW1 for this attempted connection is "connected," which prevents completion in SW1 for the D4 to HBA#2 attempted connection.

[0020]   The D2 to HBA#1 connection is "connected" in SW2 but "pending" in SW1. The "pending" state in SW1 is a result of the "connected" state in SW1 of the HBA#1 to D3 attempted connection. Turning again to the HBA#2 to D1 attempted connection, as discussed above, the state of SW1 for this attempted connection is "connected." However, the "connected" state in SW2 of the D2 to HBA#1 attempted connection prevents completion in SW2 of the HBA#2 to D1 attempted connection. Thus, the state of SW2 for the HBA#2 to D1 attempted connection is "pending."

[0021]   As a result, there is deadlock on SW1, since all pending OPN's are locked up on all four Trunks waiting for each other to complete a connection, which cannot happen.

[0022]   One way to address the lockup condition is to detect it and clear it. A way of detecting the above condition would be -- in SW#1 -- to detect the following state:

a)   Pending OPN on either Trunk#1 or Trunk#2;

b)   Pending OPN on either Trunk#3 or Trunk#4;

c)   For each of the pending OPN's described in a) and b), the destination is already in a connection on SW1.

d)   The pending OPN's and connections in c) have not seen a RRDY or SOF in a specified amount of time.

[0023]   To clarify d), seeing a RRDY or SOF would indicate that communication is occurring over the attempted connection. Similarly, the absence of an RRDY or SOF indicates that communication is not occurring over the attempted connection. One way to clear the lockup condition, when detected, is to CLS one of the pending OPN's to allow the other pending OPN's to connect to their destinations.

[0024]   A second example of a lockup situation is illustrated in Figure 3.

[0025]   Another term is explained:

a)   Close counter: Assume that D2 wants to connect to HBA#2, but the connection is closed down by a higher priority OPN from String#1. Since D2 was unable to send a frame, the Close counter gets incremented. D2 will retry and send another OPN in which it also gets closed down without sending a frame and the Close counter gets incremented again. If the close counter reaches a predefined max value, then D2 gets highest priority on the connection attempt. For more discussion of the Close counter, reference is made to the '862 application (cited above in the RELATED APPLICATIONS section).

[0026]   We now turn to another example of a lockup condition. In the Figure 3 configuration, there are 3 switches connected together by string ports. HBA#1 wants to connect to D1 and HBA #2 wants to connect to D3. D2 has a pending OPN on its port, and D4 has a pending OPN on its port. The pending OPN's of both D2 and D4 have their Close counters at max value.

[0027]   In one scenario, HBA#1 then ARB's for String#1 and wins, so HBA#1 sends an OPN that reaches SW#2 on String#1. The OPN sent by HBA#1 is now pending, because the pending OPN from D2 cannot be closed down, due to the Close counter for D2 reaching max value. HBA#2 now ARB's for String#2 and wins, so HBA#1 sends an OPN that reaches SW#3 on String #2. The OPN sent by HBA#2 is now pending because the pending OPN

from D4 cannot be closed due to the Close counter for D4 reaching max value. There is now a lockup condition in which both String #1 and String #2 have been locked up. Traffic is stopped.

[0028]   To address the stoppage of traffic, the lockup condition may be detected and cleared. A way of detecting the above condition includes detecting the following states with respect to SW#2:

a)   Pending OPN on one of String#1 and String#2;

b)   A connection on the other of String#1 and String#2, that has not seen a RRDY or a SOF in a specified amount of time;

c)   A pending OPN on a port other than on String#1 or String#2, and the destination of the OPN is either String#1 or String#2 and

d)   The pending OPN described in c) has its Close counter at max value.

[0029]   If the above conditions are met, then there is a lockup condition and one of the pending OPN's can be CLS'ed down to allow the other pending OPN's to connect to their destinations.

[0030]   FIG. 4 illustrates, in accordance with a broad aspect, a method to solve lockup conditions while trunking in Fibre Channel switched arbitrated loop systems. The method operates within a particular switch of, for example, the FIG. 2 or FIG. 3 configuration. At step 402, a particular combination of pending OPN conditions is detected, indication a lockup condition. At step 404, at least one of the detected pending OPN conditions is closed, which alleviates the lockup condition.